



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/692,290	10/23/2003	Anssi Ramo	944-003.182	7122

4955 7590 11/24/2006

WARE FRESSOLA VAN DER SLUYS &
ADOLPHSON, LLP
BRADFORD GREEN, BUILDING 5
755 MAIN STREET, P O BOX 224
MONROE, CT 06468

EXAMINER

PIERRE, MYRIAM

ART UNIT PAPER NUMBER

2626

DATE MAILED: 11/24/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No. 10/692,290	Applicant(s) RAMO ET AL.	
	Examiner Myriam Pierre	Art Unit 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 11 September 2006.
- 2a) ☒ This action is FINAL. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1 and 3-48 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1,3-48 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. _____.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

1. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

Response to Arguments

2. Applicant's arguments filed 09/11/2006 have been fully considered but they are not persuasive.

Applicant argues that Gersho et al. (154), herein referred as Gersho does not partition base on class, this argument is not persuasive, Gersho et al. incorporates analysis-by-synthesis coder which is a parametric coding which encompasses the claimed invention of segmenting, coding and decoding of audio signals, the act of parametric coding is based on a classification of the signal based on fixed criteria of the signal (col. 3 lines 1-15, 60-67 and col. 4 lines 1-6).

Applicant argues that Gersho's does not teach partitioning the speech signals into segments based on the energy characteristics of the speech signal and that the coder is not a parametric coder as disclosed in the present invention. This argument is not persuasive. Gersho teach partitioning the speech signals into segments based on the energy characteristics of the speech signal. When the energy peak is determined, the speech signal is already portioned into frames. The energy is used to determine if the frame is voice or unvoiced, and Gersho have two classifying systems in one, the first one classifies based on if a frame is voiced or not unvoiced (therefore, the energy of the frame had to be calculated), next, a second classifier is used for classifying a not unvoiced frame as being one of voiced frame or a transition frame, col. 4 lines

Art Unit: 2626

50-56 and col. 16 lines 16-23. Gersho do disclose or even suggest partitioning the speech signals into segments based on the energy characteristics of the speech signal.

CELP-type encoding, an example of Analysis-by-Synthesis coder is parametric coding which encompasses the claimed invention of segmenting, coding and decoding of audio signals for data transmission used in wireless systems (col. 3 lines 1-15, 60-67 and col. 4 lines 1-6).

Applicant argues that Gersho do not disclose or even suggest segmenting the audio signal into a plurality of segments based on the audio characteristics of the audio signal or that classifying the frames after the samples of speech signal are partitioned into frames. This argument is not persuasive. According to the abstract, "the speech is partitioned into frames and sub-frames. Performance is enhanced by coding the important segments of the excitation more accurately. Therefore, Gersho do teach segmenting the audio signal into a plurality of segments based on the audio characteristics of the audio signal. And further argues that that Gersho do not teach assigning voicing values to the voice characteristics and the segmenting is carried out based on the assigned voicing values. This argument is not persuasive. Gersho do teach assigning voice values using a two bit encoder to identify the class, firstly, frames are classified as strongly periodic, weakly periodic, erratic and unvoiced, secondly, the frames are sent to the frame classifier using two bit coding scheme, lastly, are segmented as the voiced frames are divided into three more sub-frames, col. 14 lines 7-14 and col. 9 lines 49-67. Gersho do teach assigning voicing values to the voice characteristics and the segmenting is carried out based on the assigned voicing values.

Applicant argues that Gersho do not disclose or suggest that the partitioning of the samples into frames is based on which part of the excitation is encoded and partitioning of the

Art Unit: 2626

samples into frames is based on target accuracy. This argument is not persuasive. Gersho teach partitioning which necessarily would include a target accuracy, col. 7 lines 18-26 and Fig. 2. And also discloses the partitioning of the samples into frames is based on target accuracy as in the energy, part of the excitation, is used to determine if the frame is voice or unvoiced, and Gersho have two classifying systems in one, the first one classifies based on if a frame is voiced or not unvoiced (therefore, the energy of the frame had to be calculated), next, a second classifier is used for classifying a not unvoiced frame as being one of voiced frame or a transition frame, col. 4 lines 50-56 and col. 16 lines 16-23. Gersho do disclose or suggest that the partitioning of the samples into frames is based on which part of the excitation is encoded.

Applicant argues that Gersho only discloses which part of the excitation is used in encoding. That has nothing to do with providing a linear pitch representation in some of the segments. This argument is not persuasive. Gersho also discloses relaxation CELP which ensure that the input signal conforms to a simplified (linear) pitch contour, col. 2 lines 14-18.

Applicant argues that the speech and the update prediction parameters are not audio data indicative of the parameters as received in a decoder as claimed. This argument is not persuasive. Gersho do teach assigning voice values using a two bit encoder to identify the class, firstly, frames are classified as strongly periodic, weakly periodic, erratic and unvoiced, secondly, the frames are sent to the frame classifier(parameters) using two bit coding scheme, lastly, are segmented as the voiced frames are divided into three more sub-frames, col. 14 lines 7-14 and col. 9 lines 49-67. Gersho do teach update prediction parameters are audio data indicative of the parameters as received in a decoder as claimed.

Claim Rejections - 35 USC § 102

3. Claims 1, 3-14, 19-38, 40-4-48 are rejected under 35 U.S.C. 102 (b) as being anticipated by Gersho et al. (6,311,154).

As to claim 1, Gersho et al. teach

segmenting {partitioning or classifying} the audio signal {speech} into a plurality of segments {frames} (partitioning samples of a speech signal into frames, col. 4, lines 25-27) based on the audio characteristics {classes} of the audio signal (classifying the speech signal in each from into one of a plurality of classes, col. 4, lines 25-27); and

encoding the segments {frames} with different encoding settings {excitation} (encoding an excitation for the frame using one of a plurality of excitation coding...selected according to the class of the frame, col. 4, lines 30-33).

As to claim 3, Gersho et al. teach

characteristics {classes/classifying} include voicing characteristics {voice} in said segments {frames} of the audio signal {speech signal} (classifying the speech signal in each frame into classes, classes include voice frame, col. 4, lines 25-27 & 35).

As to claim 4, Gersho et al. teach

characteristics {identifying} include energy characteristics {presence of energy} in said segments {window} of the audio signal {residual signal} (identifying the location of a window, identifying considers the presence of energy in the residual signal, col. 4, lines 65-67).

As to claim 5, Gersho et al. teach

Characteristics {positioning} include pitch characteristics {function of the pitch} in said segments {frames} of the audio signal (positioning the window at a location that is a function of a pitch of the frame, col. 4, lines 59-61).

As to claim 6, Gersho et al. teach

segmenting {partitioning} is carried out concurrently {classifying and encoding} to said encoding {coding} (partitioning samples of speech, classifying speech signals into classes, coding a speech signal, col. 4, lines 24-25. The classifying and encoding process may be done concurrently).

As to claim 7, Gersho et al. teach segmenting is carried out before said encoding (partitioning samples of speech, classifying speech signals into classes, coding a speech signal, col. 4, lines 24-25, thus the classifying or segmenting is done before coding).

As to claim 8, Gersho et al. teach

plurality of voicing values {voice or unvoiced} are assigned to the voicing characteristics of the audio signal in said segments, and wherein said segmenting {partitioning} is carried out based on the assigned voicing values (classifying a frame is being one of an unvoiced or voiced, col. 4, lines 52-53).

As to claim 9, Gersho et al. teach

Art Unit: 2626

a value designated {classifying} to a voiced speech signal and another value designated to an unvoiced signal (classifying a frame is being one of an unvoiced or voiced, col. 4, lines 51-52).

As to claim 10 Gersho et al. teach

A value designated {classifier} to a transitional stage between the voice and unvoiced {transitional} signals {frame} (classifier for classifying a transition frame, col. 4, lines 52-55).

As to claim 11, Gersho et al. teach

a value designated {(m)=1} to an inactive period {silent frame} in the audio signal {speech} (If (m)=1, then the current frame is declared a silent frame, col. 15, lines 7-8 & 35-37).

As to claim 12, Gersho et al. teach

selecting a quantization mode for said encoding in order to improve the bit allocation and to reduce the parameter update rate, wherein the segmenting is carried out based on the selected quantization mode (col. 3 lines 45-49; Fig. 5 and col. 11 lines 4-16; col. 4, lines 36-37, col. 15, lines 35-36 & col. 9, lines 63-65).

As to claim 13, Gersho et al. teach

segmenting is carried out based on target accuracy in reconstruction of the audio signal, wherein the target accuracy is selected based on distortion criteria comparing up-sampled quantized values (transmitted samples) and modified parameters (col. 9, lines 63-65 and col. 3 lines 45-49).

As to claim 14, Gersho et al. teach
segmenting is carried out for providing a linear pitch representation in at least some of
said segments (col. 9, lines 63-65; col. 3 lines 45-49 and col. 4 lines 50-62).

As to claim 19 and 27, Gersho et al. (154) teach
an input for receiving audio data indicative of the parameters in the adjusted
representation (input applied to element 14, Fig. 3).

and a module responsive to the audio data for generating the audio signal based on the
adjusted signals and the characteristics of the audio signal (Fig. 3. One would necessarily need a
module to respond to an adjusted audio signal/characteristics of audio signals).

At the time of the invention, it would have been obvious to one of ordinary skill in to use
a decoder in order to reverse the encoding data for further processing, such as modulating or
storing the audio signal.

As to claim 20 and 28, Gersho et al. (154) does not teach recording parameters.

At the time of the invention, it would have been obvious to one of ordinary skill in the art
to record audio parameters in order to update the audio data for storage and retrieval.

As to claim 21 and 29,

Gersho et al. (154) teach.

Art Unit: 2626

the audio data is transmitted through a communication channel and wherein the input of the decoder is operatively connected to the communication channel for receiving the audio data (digital communications, col. 1, line 1 and Fig. 3).

As to claim 22, Gersho et al. (154) teach,

an input for receiving audio data indicative of the characteristics (encoder, Fig. 1, element 82); and

an adjustment module for adjusting a parameter based on the characteristics of the audio signal for providing an adjusted representation of a parameter, wherein said adjusting comprises segmenting the audio signal into a plurality of segments based on the characteristics of the audio signals and encoding the segments based on one or more of a plurality of encoding settings (LP coding, modified residual, adjusts frames, Abstract and Fig. 9; col. 8 lines 54-63).

As to claim 23, Gersho et al. (154) teach,

a quantization module responsive to the adjusted representation for coding the parameters in the adjusted representation (Fig. 9).

As to claim 24, Gersho et al. (154) teach,

an output end operatively connected to a storage medium for providing data indicative of the coded parameters in the adjusted representation (stored as vectors in a codebook, col. 1, lines 64-65).

As to claim 25, Gersho et al. (154) teach,
output end, operatively connected to a communication channel for providing signals indicative of the coded parameters in the adjusted representation to the communication channel for transmission (Fig. 8; a coder which necessarily has an output and ability to represent the adjusted audio parameters).

As to claim 26, Gersho et al. (154) teach,
a code for determining the characteristics of the audio signal (LP coding, col. 8 lines 54-63)

a code for adjustment the parameter based on the characteristics of the audio signal for providing an adjusted representation of the parameter, wherein said adjusting comprises the steps of segmenting the audio signal into a plurality of segments based on the characteristics of the audio signal and encoding the segments based on one or more of a plurality of encoding settings (LP coding, modified residual, adjusts frames, Abstract and Fig. 9; col. 8 lines 54-63).

As to claim 30, Gersho et al. (154) teach
a mobile terminal (mobile base station, col. 6, lines 17-18).

As to claim 31, Gersho et al. (154) teach,
Implementing in a cell phone system which necessarily has both base station and mobile station adapted to communicating with the base stations (col. 6, lines 33-36).

a decoder for use in parametric audio coding for generating a synthesized audio signal indicative of an audio signal having audio characteristics, wherein the audio signal is coded in a coding step into a plurality of parameters at a data rate and the encoding step is adjusted based on the characteristics of the audio signal for providing an adjusted representation of the parameters, wherein the said adjusting comprises the steps of segmenting the audio signal into a plurality of segments based on the characteristics of the audio signal and encoding the segments based on one or more of a plurality of encoding settings (Figs 1, 4-5 , LP coding, modified residual, adjusts frames, Abstract and Fig. 9; col. 8 lines 54-63).

an input for receiving audio data indicative of the parameters in the adjusted representation from at least one of the base stations for providing the audio data to the decoder, so as to allow the decoder to generate the synthesized audio signal based on the adjusted representation (Figs 1, 4-5, col. 3 lines 1-15).

As to claim 32, Gersho et al. (154) teach,

an input for receiving audio data indicative of end points defining a plurality of sub-segments, wherein the audio signal is encoded for providing parameters indicative of the audio signal, the parameters including pitch contour data containing a plurality of pitch values representative of an audio segment in time, and wherein the pitch contour data in the audio segment in time is approximated by a plurality of consecutive sub-segments in the audio segment, and wherein the end points include a first end point and a second end point for defining each of said sub-segments (decoder, col. 6 lines 8-11 and Fig. 1); and

a reconstruction module for reconstructing the audio segment based on the received audio data (Fig. 9; col. 6 lines 8-11).

As to claim 33, Gersho et al. (154) teach, encoding settings inherently include bit allocation (col. 3 lines 45-49), quantization accuracy (Fig. 5 and col. 11 lines 4-16), quantization method (col. 11 lines 4-16) and parameter update rate (col. 3 lines 31-44 and 56-60).

As to claim 34, Gersho et al. (154) teach, the audio signal contains sinusoidal components (col. 3 lines 25-29, analysis windows made equal becomes sine) and said parameters include frequency values (Fig. 1 element 68), amplitude values (col. 3 lines 51-55) and phase values indicative of the sinusoidal components (Fig. 1 element 76 and col. 3 lines 25-29).

As to claim 35, Gersho et al. (154) teach, the parameters includes pitch (col. 4 line 60), voicing f(Fig. 9 element 42c), amplitude (col. 3 lines 51-55) and energy of the audio signal (col. 3 lines 42-44).

As to claim 36, Gersho et al. (154) teach, the parameters include pitch contour data (col. 4 line 60-61) containing a plurality of pitch values inherently representative of an audio segment in time (col. 4 lines 59-63 and col. 2 lines 51-64).

Art Unit: 2626

As to claim 37, Gersho et al. (154) teach, encoding settings inherently include bit allocation (col. 3 lines 45-49), quantization accuracy (Fig. 5 and col. 11 lines 4-16), quantization method (col. 11 lines 4-16) and parameter update rate (col. 3 lines 31-44 and 56-60, Fig. 4, 8-9 and 14).

As to claim 38, Gersho et al. (154) teach, encoding settings inherently include bit allocation (col. 3 lines 45-49), quantization accuracy (Fig. 5 and col. 11 lines 4-16), quantization method (col. 11 lines 4-16) and parameter update rate (col. 3 lines 31-44 and 56-60, and col. 3 lines 1-15).

As to claim 40, Gersho et al. (154) teach, encoding settings inherently include bit allocation (col. 3 lines 45-49), quantization accuracy (Fig. 5 and col. 11 lines 4-16), quantization method (col. 11 lines 4-16) and parameter update rate (col. 3 lines 31-44 and 56-60, col. 6 lines 8-11).

As to claim 41, Gersho et al. (154) teach, wherein the audio signal comprises a plurality of frames and the audio signal in each frame has a waveform and wherein the further audio signal is produced in the decoding stage independently of the waveform (col. 14 lines 8-14; col. 13 lines 62-67 and col. 14 lines 1-7).

As to claim 42, which depends on claim 1, Gersho et al. (154) teach

Art Unit: 2626

wherein each segment has a segment length and wherein the segment length of at least one segment is different from the segment length of at least one other segment (col. 14 lines 8-14; col. 13 lines 62-67 and col. 14 lines 1-7).

As to claim 43, which depends on claim 19, Gersho et al. (154) teach wherein the audio signal comprises a plurality of frames and the audio signal in each frame has a waveform and wherein the module generates the further audio signal independently of the waveform (col. 14 lines 8-14; col. 13 lines 62-67 and col. 14 lines 1-7).

As to claim 44, which depends on claim 19, Gersho et al. (154) teach wherein the segments comprise segments of different segment lengths (col. 14 lines 8-14).

As to claim 45, which depends on claim 22, Gersho et al. (154) teach wherein the segments comprise segments of different segment lengths (col. 14 lines 8-14).

As to claim 46, which depends on claim 26, Gersho et al. (154) teach wherein the segments comprise segments of different segment lengths (col. 14 lines 8-14).

As to claim 47, which depends on claim 31, Gersho et al. (154) teach

Art Unit: 2626

wherein the segments comprise segments of different segment lengths (col. 14 lines 8-14).

As to claim 48, which depends on claim 32, Gersho et al. (154) teach wherein the segments comprise segments of different segment lengths (col. 14 lines 8-14).

Claim Rejections - 35 USC § 103

4. Claims 15-18 are rejected under 35 U.S.C. 103(a) as being unpatentable over Gersho et al. (6,311,154) as applied to claim 1 above, and further in view of Gersho (IEEE-96).

As to claim 15, Gersho et al. (154) teach

Forming a parameter signal based on the audio signal data {speech} data having a first number (speech, col. 15, lines 1-20).

Gersho et al. (154) do not explicitly teach down-sampling.

However, Gersho (IEEE-96) teach down-sampling the parameter signal to a second number of a signal for providing a further parameter signal, wherein the second number is necessarily smaller than the first number (down-sampling, page 905, right col., paragraph 1. down-sampling necessarily having a smaller number than the first; second number would necessarily be smaller than the first because it is being counted backwards or decremented, starting with the last number first, in a down-sampling

Art Unit: 2626

process or when up-sampling wherein the third sample is greater than the second, which is another example of the down-sample in reverse).

At the time of the invention, it would have been obvious to one of ordinary skill in the art to down-sample the encoded speech signal, in order to reduce sampling rate, thus providing a large complexity reduction, as taught by Gersho (IEEE-96), page 905, right col., paragraph 1.

Neither Gersho et al. (154) nor Gersho (IEEE-96) explicitly teach up-sampling.

At the time of the invention, it would have been obvious to one of ordinary skill in the art to up-sample the encoded speech signal for decoding, and necessarily the third number is equal to or greater than the first number, in order to restore the original parameters for decoding.

As to claim 16, Gersho et al. (154) teach the third number is equal to the first number (col. 12 lines 45-51; delay estimates are necessarily going to allow a third number to be equal to the first).

As to claim 17, Gersho et al. (154) teach the signal data {speech} comprise quantized {two bits per frame} parameters (linear prediction parameters, col. 8, lines 57-58 & col. 9, line 65. Two bits per frame is used to identify the class/parameters of the speech signal, such as 00, 01, etc.).

As to claim 18, Gersho et al. (154) teach signal data comprises un-quantized {un-quantized linear prediction} parameters (parameters, col. 15, lines 14).

Conclusion

5. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire **THREE MONTHS** from the mailing date of this action. In the event a first reply is filed within **TWO MONTHS** of the mailing date of this final action and the advisory action is not mailed until after the end of the **THREE-MONTH** shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than **SIX MONTHS** from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Myriam Pierre whose telephone number is 571-272-7611. The examiner can normally be reached on Monday - Friday from 5:30 a.m. - 2:00p.m.

6. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571) 272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information as to the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR

Art Unit: 2626

system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Myriam Pierre
AU 2626
11/11/06



RICHEMOND DORVIL
SUPERVISORY PATENT EXAMINER